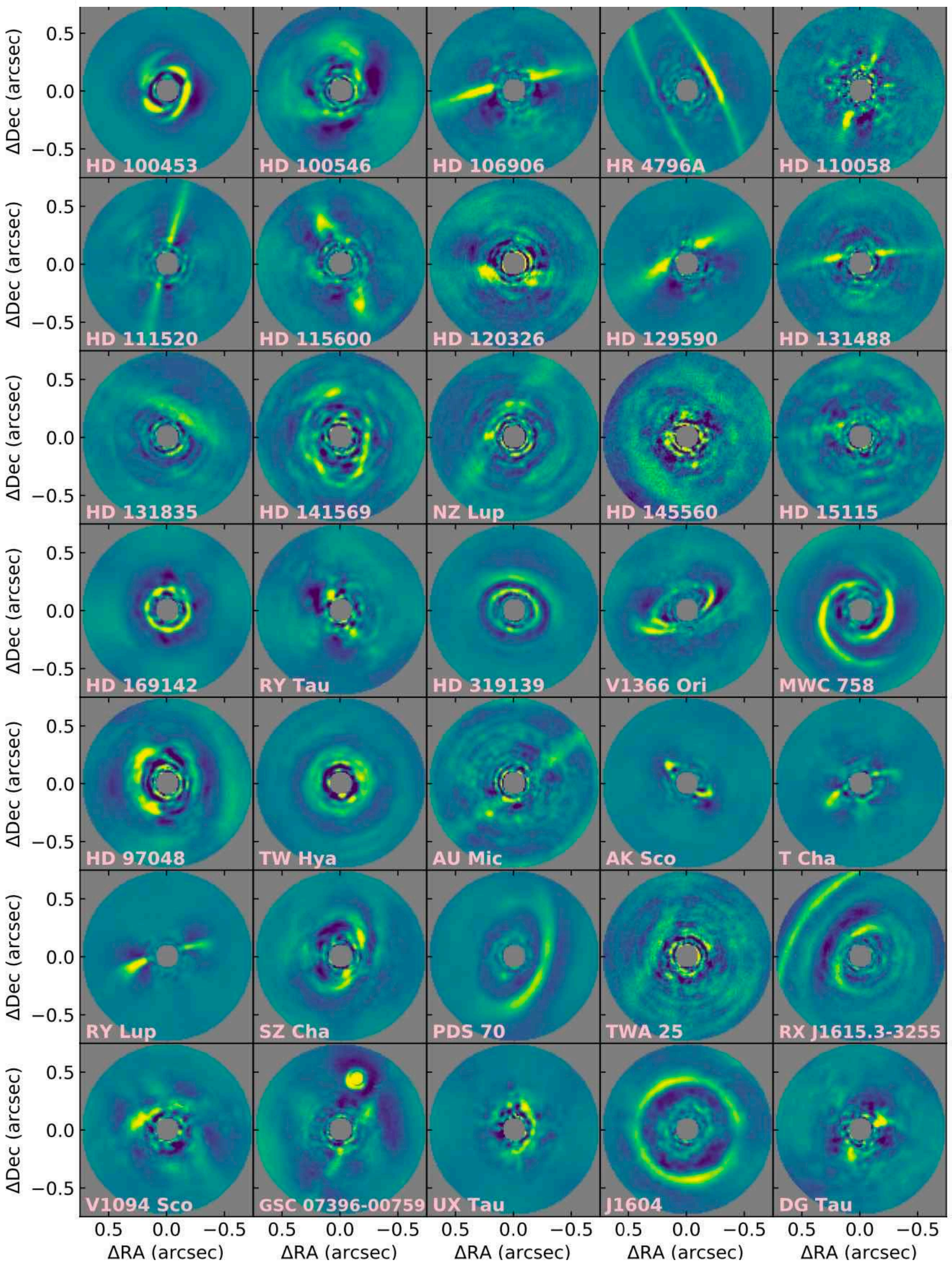
Post-processing with Reference Differential Imaging

Methodology and results

Gallery of disks post-processed in RDI using a libray of SPHERE images (Xie et al. 2022)



- History of RDI
- Practical implementation of RDI
- RDI applied to space-base and ground-based imaging
- Performance of RDI
  - In star-hopping
  - With multiple reference stars from large library (mRDI)
  - Comparison single/multiple RDI (C. Romero)

A bit of history of RDI (1/2)

Probably the first technique used historically in high-contrast scenes

- from the ground without AO (Smith & Terrile 1987)
- from the ground with AO (Beuzit et al. 1996, Mouillet et al. 1997)
- from space (Schneider et al. 1999)

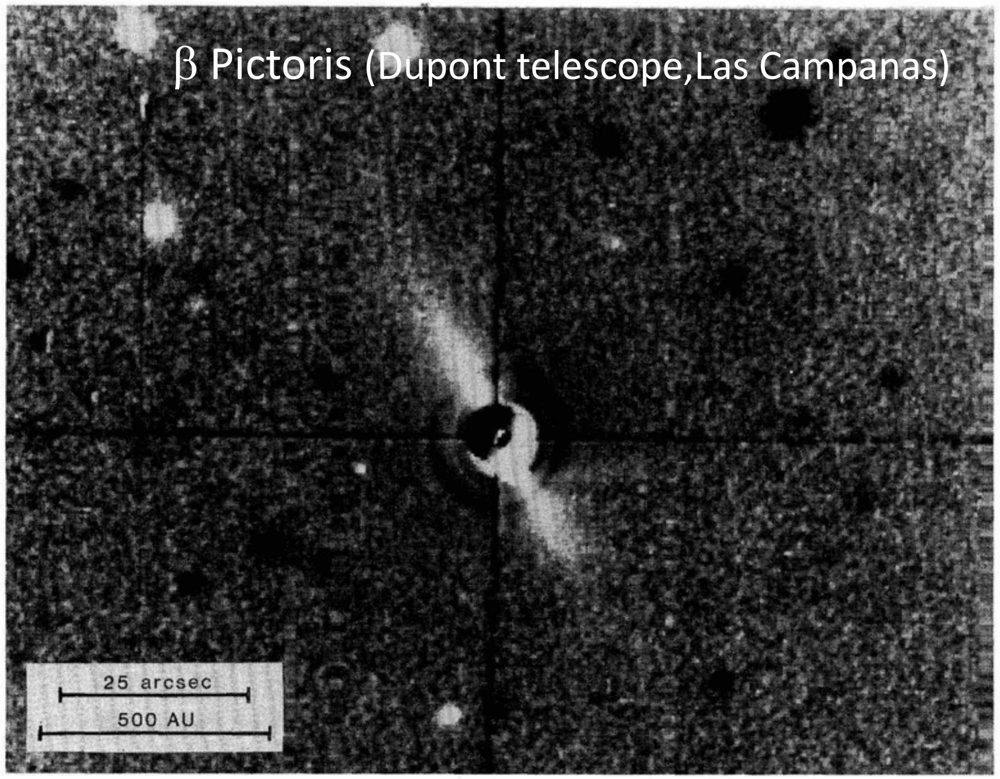
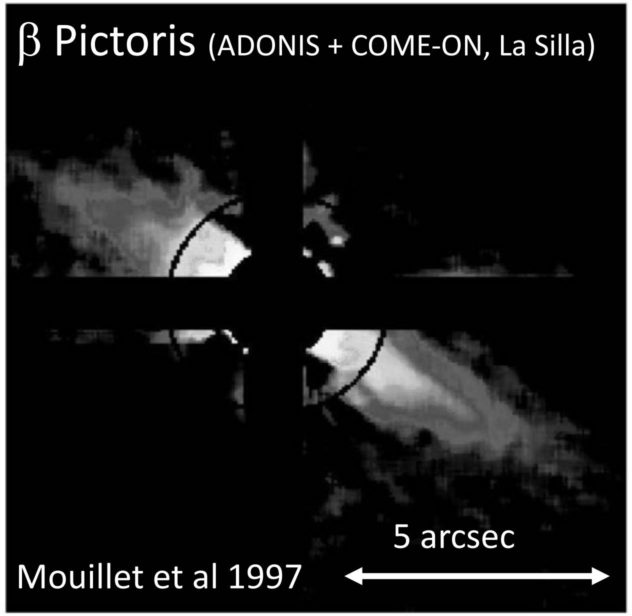
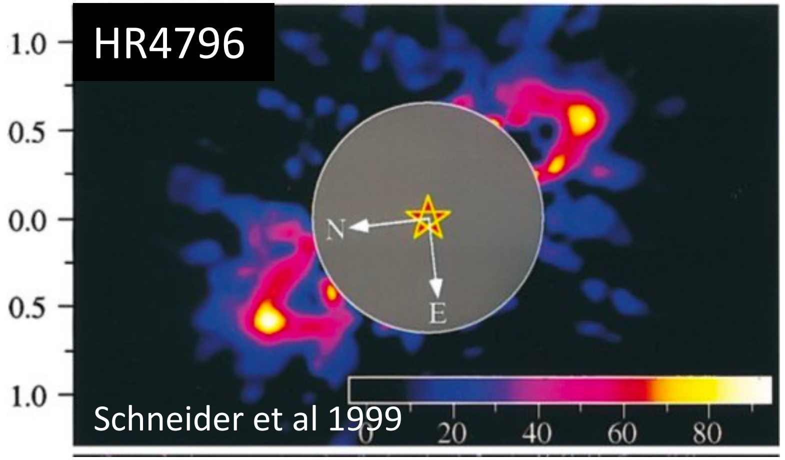


Fig. 1. Ratio image (β Pictoris divided by α Pictoris) showing the edge-on circumstellar disk extending 25 arcsec (400 AU) to the northeast and southwest of the star, which is situated behind an obscuring mask. North is at the top. The dark halo surrounding the mask is caused by imperfect balance in the ratioing process. For further explanation, see text.

Smith & Terrile, 1984 (1st detection)





A bit of history of RDI (2/2)

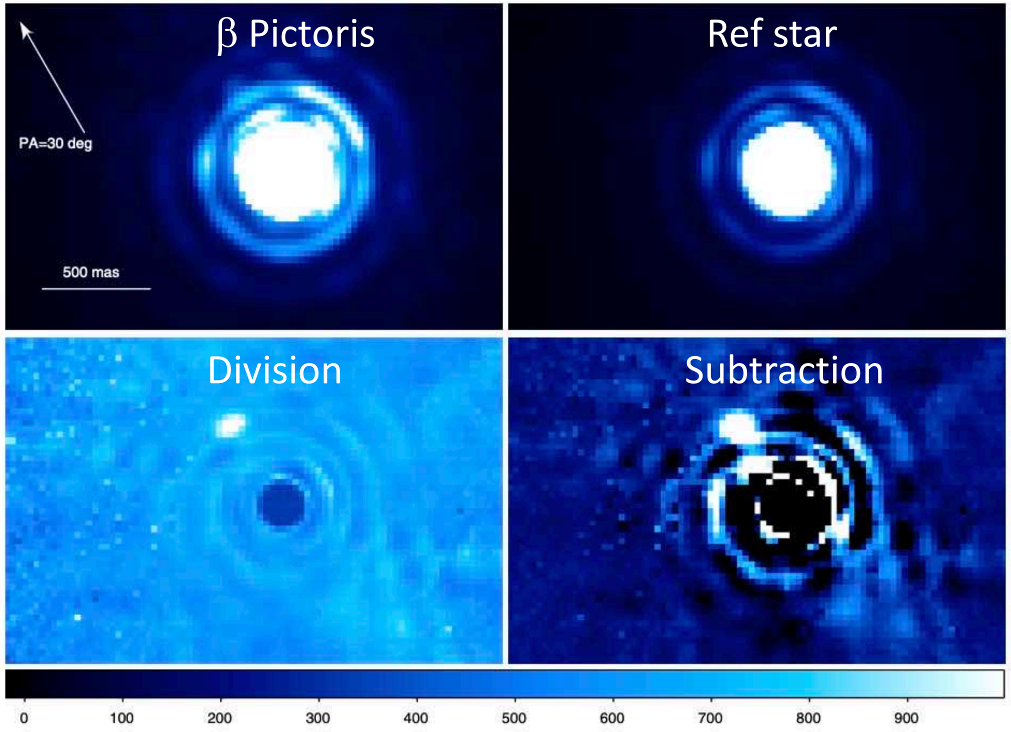


Fig. 1. β Pic and HR 2435 recentered and saturated L' images (*top left and top right*, respectively) in data set A. Below are the divided (*bottom left*) and subtracted (*bottom right*) images. North is up and east to *the left*. A candidate companion is clearly detected at a PA of $\simeq 32°$, i.e., along the NE side of the disk, at a separation of about 0.′′41 from the star.

Lagrange et al 2008

- Observations in 2003 carried out in field stabilisation → requirement to match the parallactic angles of the science and ref star to get the same pupil orientation
- Technique later superseded on the ground by observations in pupil stabilisation allowing the use of ADI
- In space, renewed interest when
  - A legacy archive of NICMOS PSF library was assembled (Laplace, Schneider et al. 2010)
  - LOCI and PCA was proposed to improve the performance of high-contrast algorithms

RDI in practice: implementation with PCA

- Description proposed in Soummer et al. 2012, Amara & Quanz 2012

ε = 0 (absence) or 1 (presence)

Science image:

$$T(n) = I_{\psi_0}(n) + \epsilon A(n),$$

unknown PSF

astrophysical signal

PSF estimation:

$$\hat{I}_{\psi_0}(n) = \sum_{k=1}^{K_{\text{klip}}} < T, Z_k^{\text{KL}} >_{\mathcal{S}} Z_k^{\text{KL}}(n),$$

KL transform of the set of reference images

Residuals :

$$F(n) = \left(I_{\psi_0}(n) - \sum_{k=1}^{K_{\text{klip}}} < I_{\psi_0}, Z_k^{\text{KL}} >_{\mathcal{S}} Z_k^{\text{KL}}(n) \right)$$

PSF residuals

$$+ \epsilon \left(A(n) - \sum_{k=1}^{K_{\text{klip}}} < A, Z_k^{\text{KL}} >_{\mathcal{S}} Z_k^{\text{KL}}(n) \right).$$

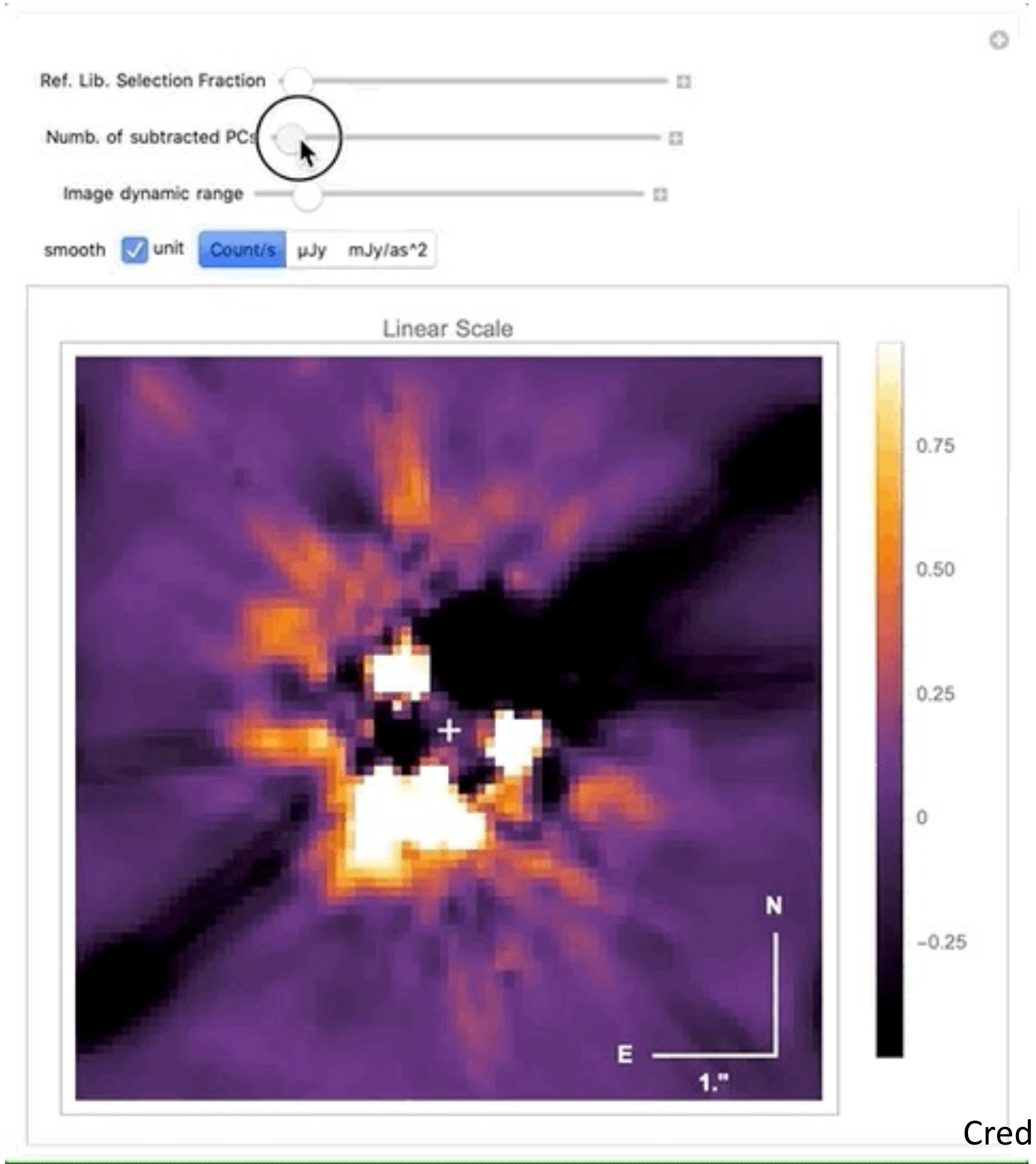
Over-subtraction (can be computed using forward modelling)

RDI in practice: implementation with PCA

Free parameters:

- Size of the reference frames library
- Number of modes subtracted (from 1 to the max number of reference frames in the library)
- Region where the subtraction is performed (full frame, annuli, small patches…)

In practice many more flavours to keep in mind: optimisation per frame/target, pre-processing: spatial/temporal mean subtraction, hybrid RDI-ADI

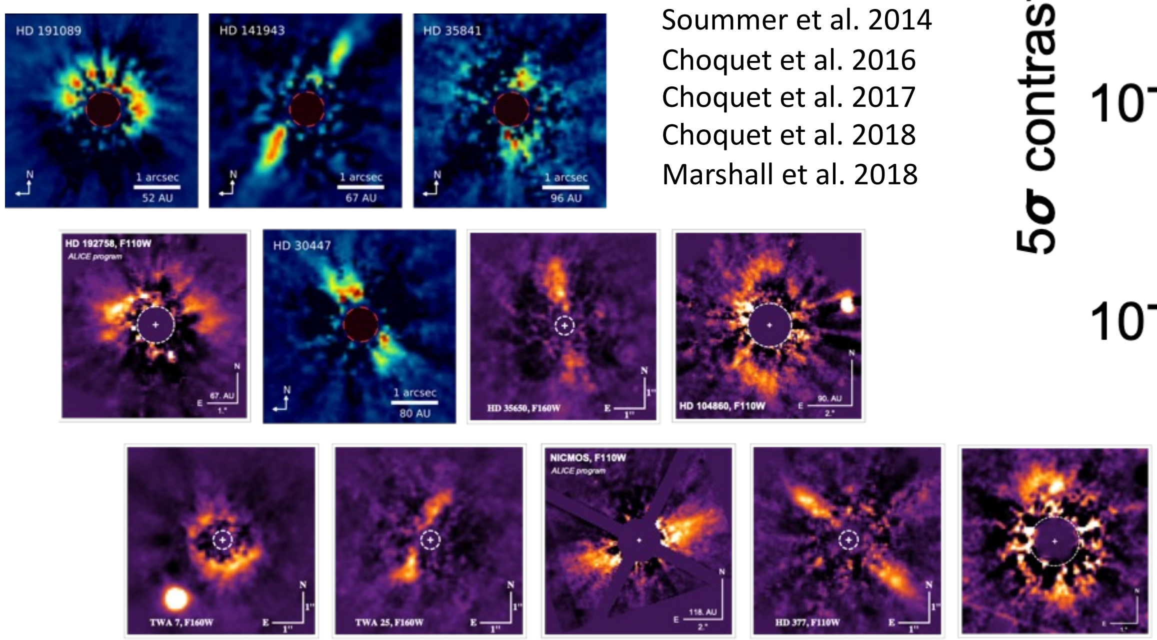


Credit: E. Choquet

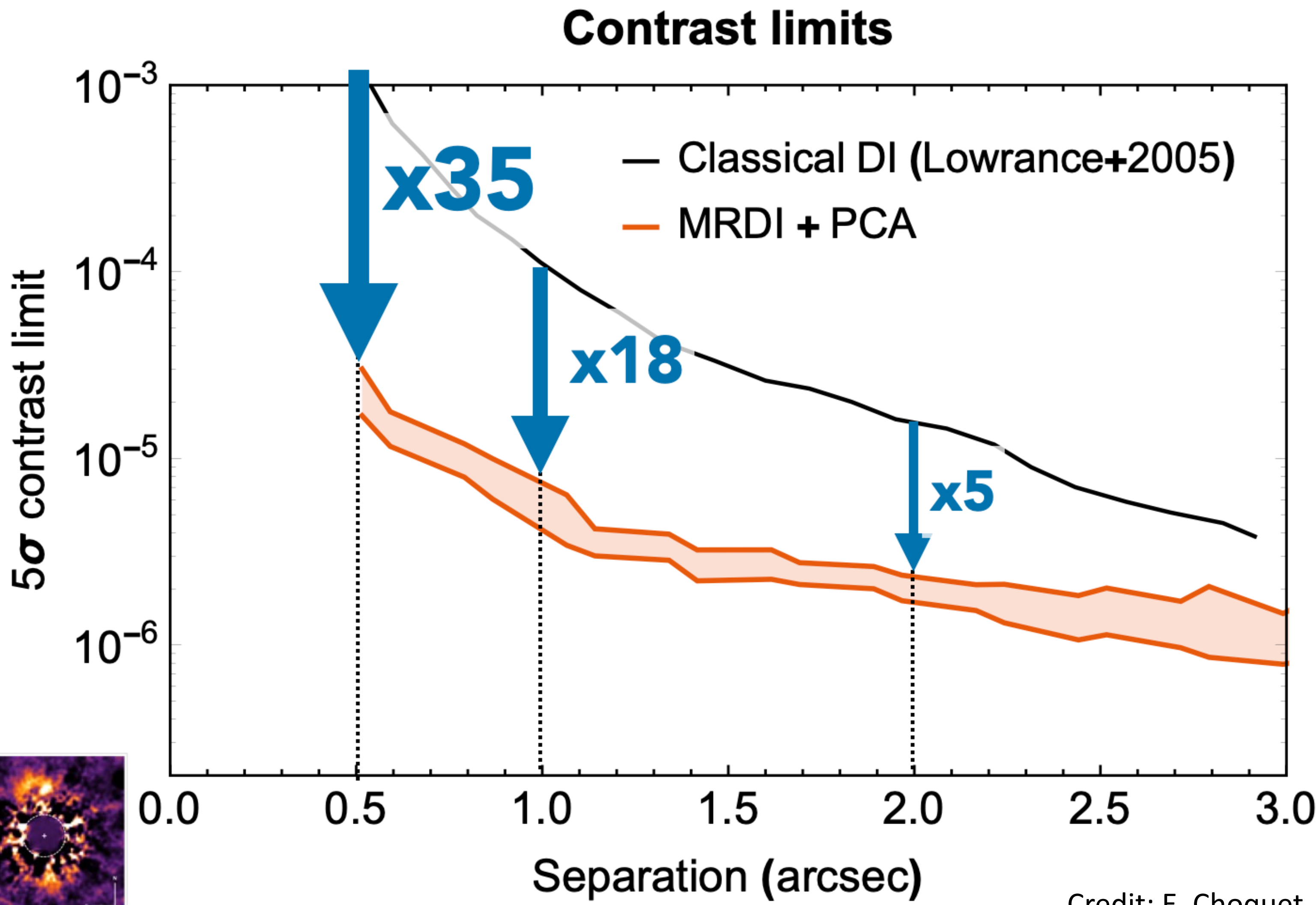
RDI from space with HST/NICMOS

Results of the ALICE program

- 12 new debris disks (~1/4 of known disks)
- redetection of HR 8799 bcd in old NICMOS data (Lafrenière et al. 2009, Soummer et al. 2011)



Soummer et al. 2014
Choquet et al. 2016
Choquet et al. 2017
Choquet et al. 2018
Marshall et al. 2018



Credit: E. Choquet

First conclusions from space-based RDI

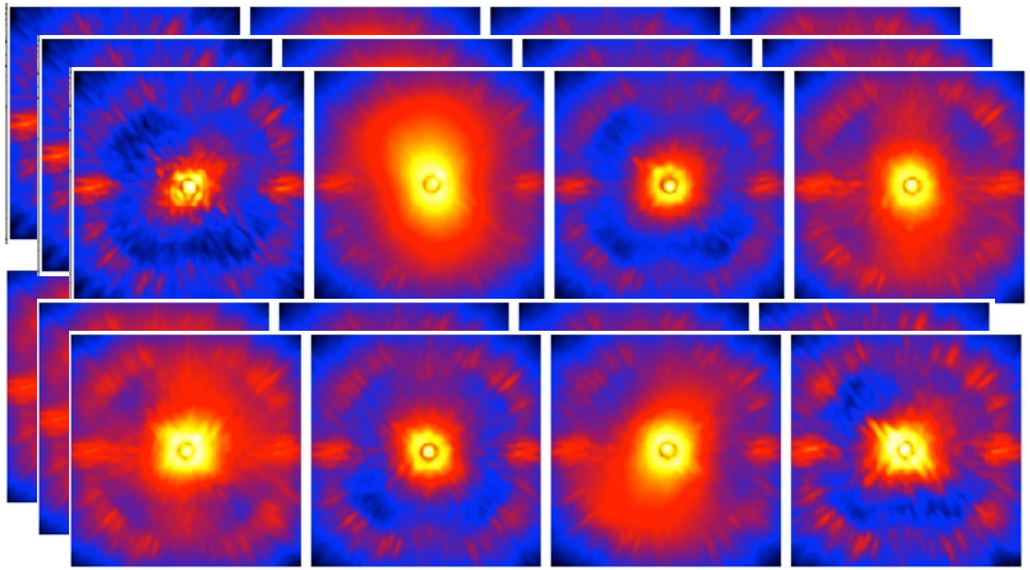
- Frame selection is critical
- Frame registration is also important
- Trade-off library size number of modes subtracted done manually

Space- vs ground-based observations

	Space	Ground-based
Star rejection	Sub-optimal Coronagraphs	Optimized coronagraphs + WFC
Variability	Thermal variations Orbit-to-orbit reproducibility	Atmospheric turbulence residuals Changing observing conditions
Obs. Strategy	Discrete Rolls Long exposures (100-1000s) Few images per dataset (~10)	Continuous ADI Short exposures (1-10s) Many images per dataset (~100)

Credit: E. Choquet

RDI with ground-based observations

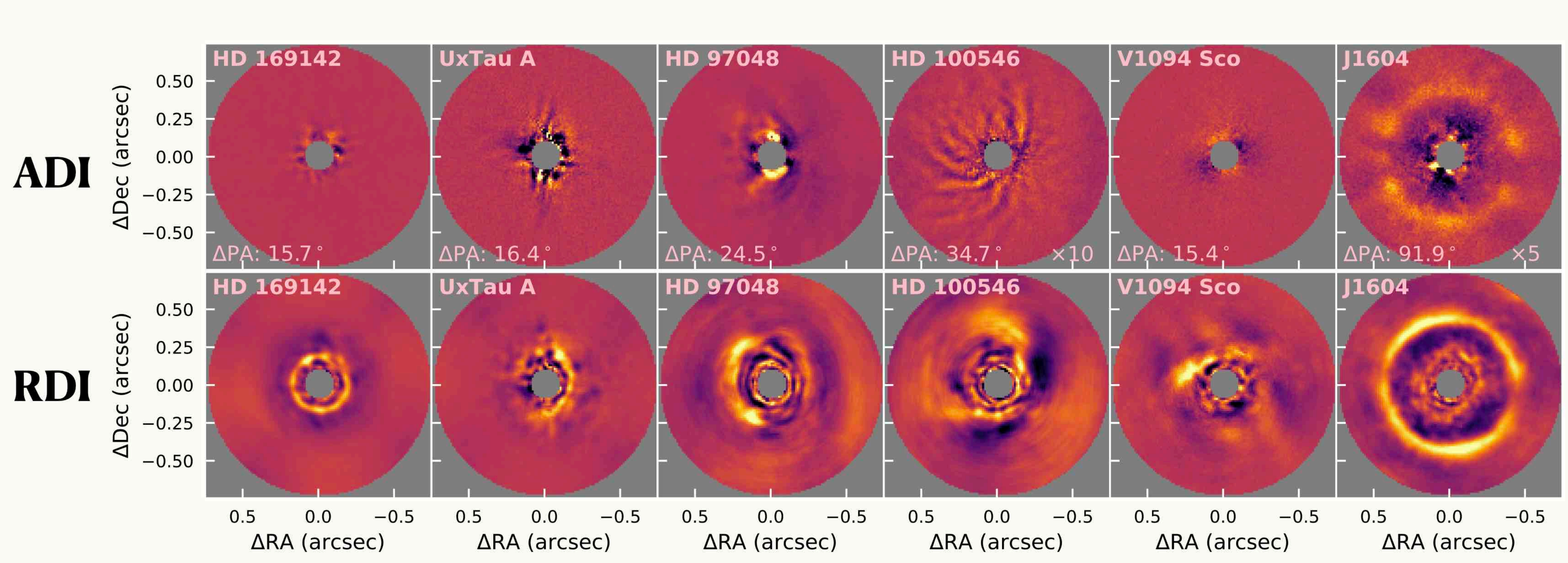


Star-hopping (Wahhaj+2021) and binary differential imaging (Rodigas+2015) showed that RDI from the ground works with single well-chosen reference frames.

What about large libraries ?

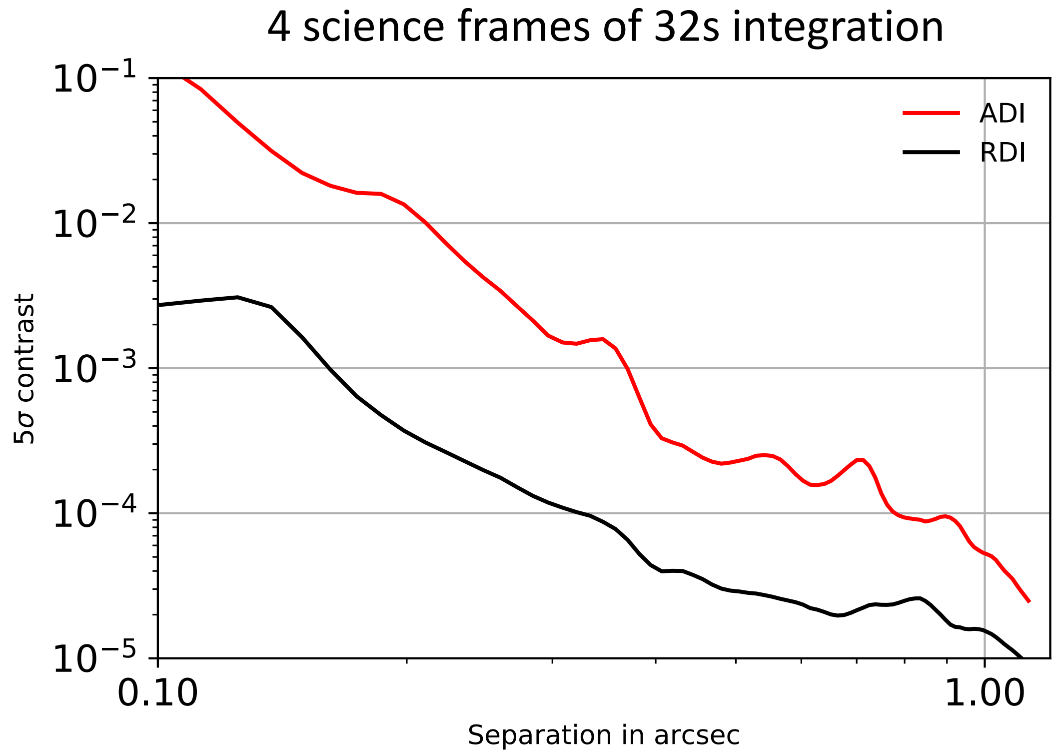
It works !

NIRC2: Xuan+2018, Ruane+2019
GPI: Draper+2016
SPHERE: Xie+2022



Gallery of protoplanetary disks from Xie et al. 2022. 4 disks are detected in intensity for the first time

Case 1: Snapshot imaging

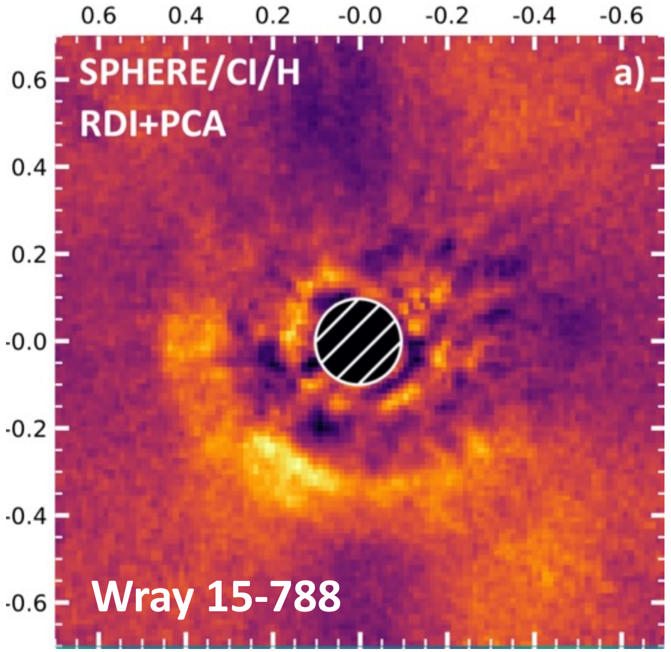


Good contrast with RDI even for short exposure time

Young Suns Exoplanets Survey (YSES, Bohn et al. 2019, 2020, 2021)

5 min on-source for each target with little field rotation (<5deg)

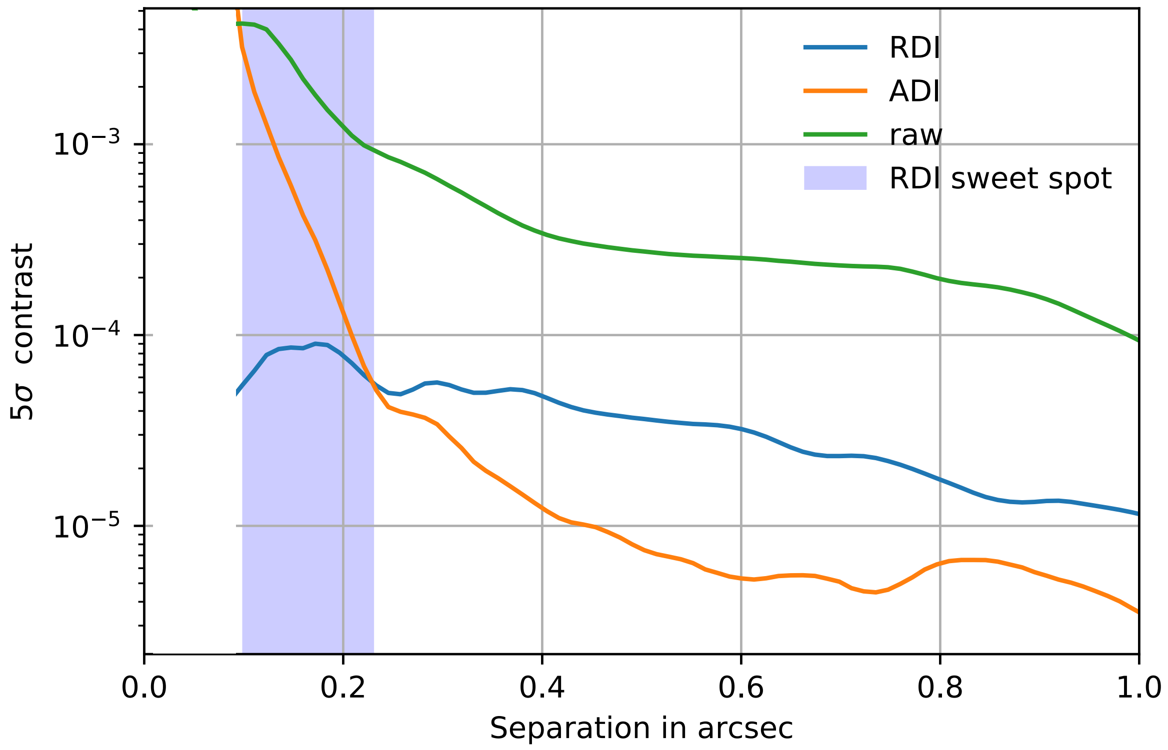
Library of ~ 200 frames on 26 stars



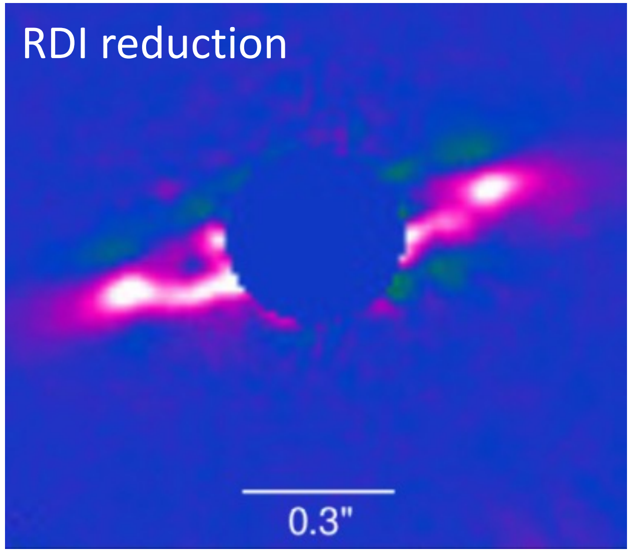
Bohn et al. 2019

Case 2: longer sequences

Example of medium-size library : SPHERE High-Angular Resolution Debris Disk Survey (SHARDDS led by J. Milli)
55 targets (~20 000 frames), 40min on source for each target with large field rotation

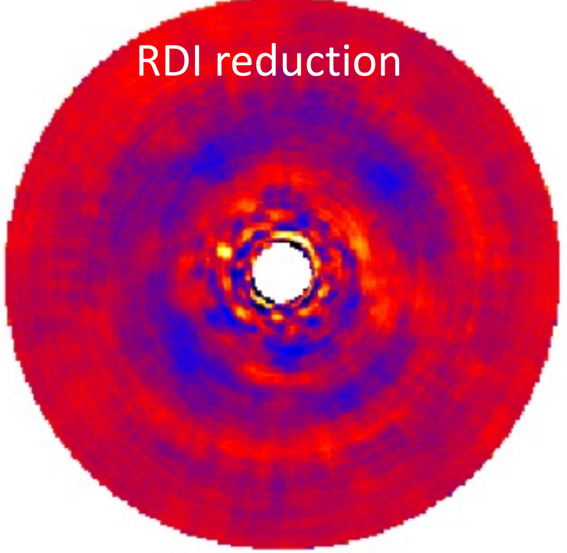


Disk of HD 114082



Wahhaj et al. 2016

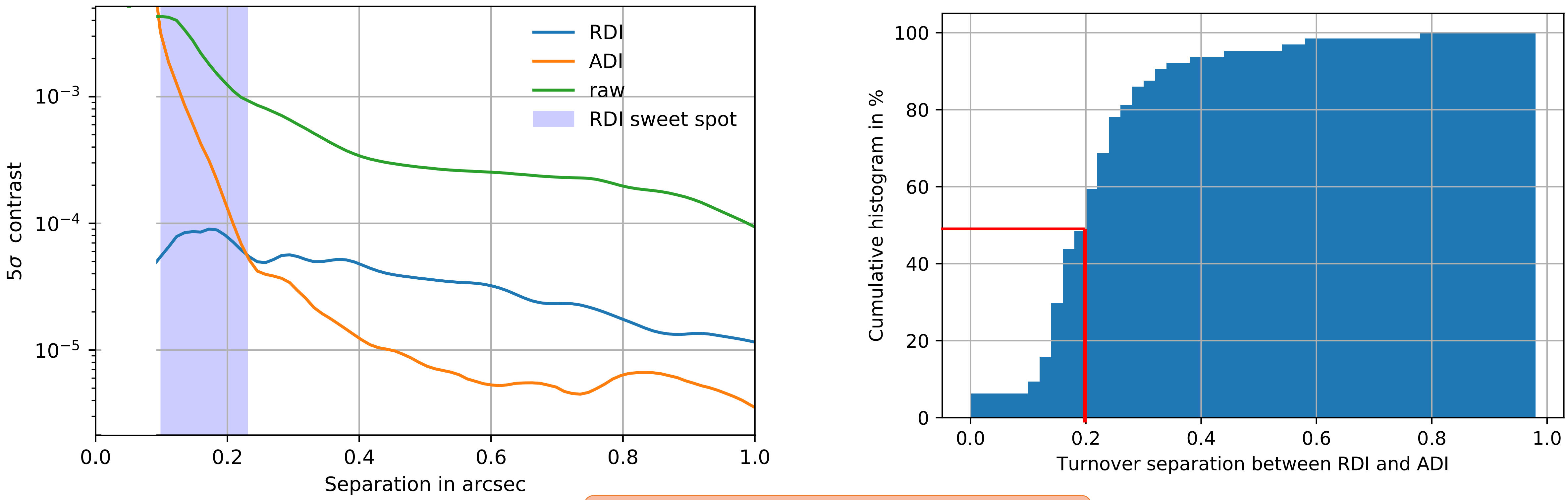
Brown dwarf HD 206893B
10^{-4} contrast at 270 mas



Milli et al. 2017
Romero et al. 2021

Case 2: deeper sequences

Example of medium-size library : SPHERE High-Angular Resolution Debris Disk Survey (SHARDDS led by J. Milli)
55 targets (~20 000 frames), 40min on source for each target with large field rotation



Results from the SHARDDS survey

Turnover point at ~200mas

Similar conclusion found in - Xie et al. 2022: RDI outperforms ADI < 400mas under 0.4-0.6 arcsec seeing conditions
- Xuane et al. 2018: turnover at 250mas on average

What about large-size libraries > 100 000 frames?

Xie et al. 2022 used the SPHERE archive in the H23 filter with ~130 000 frames.

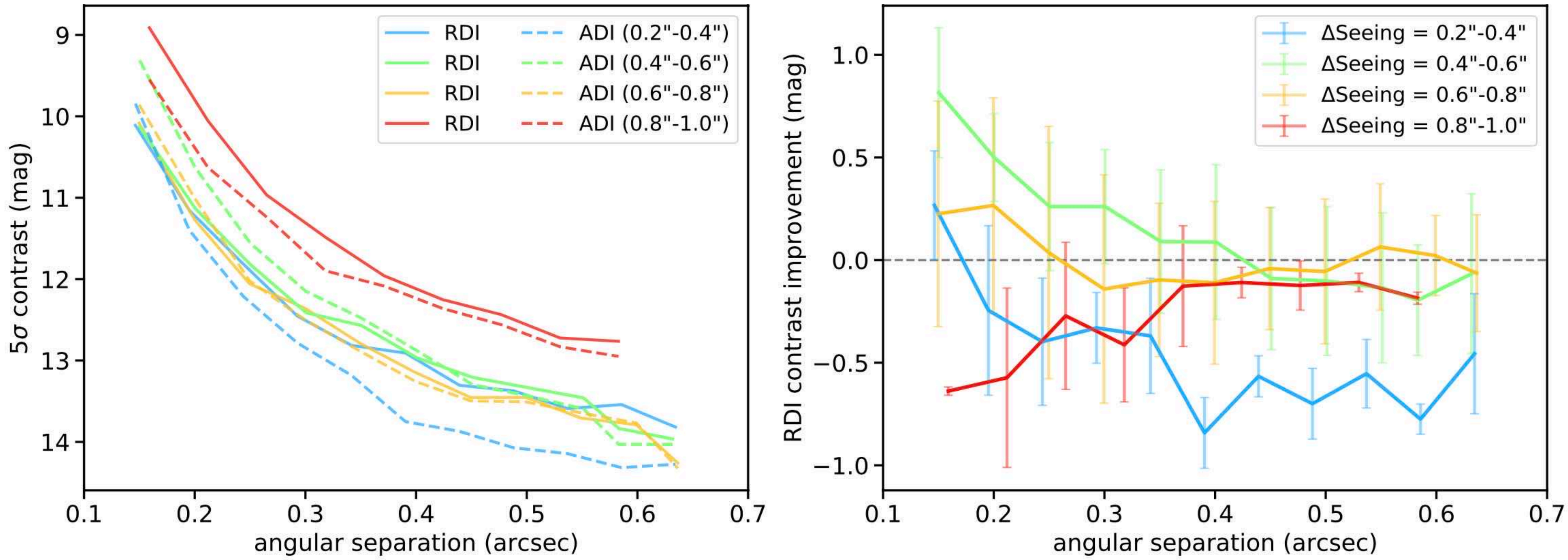
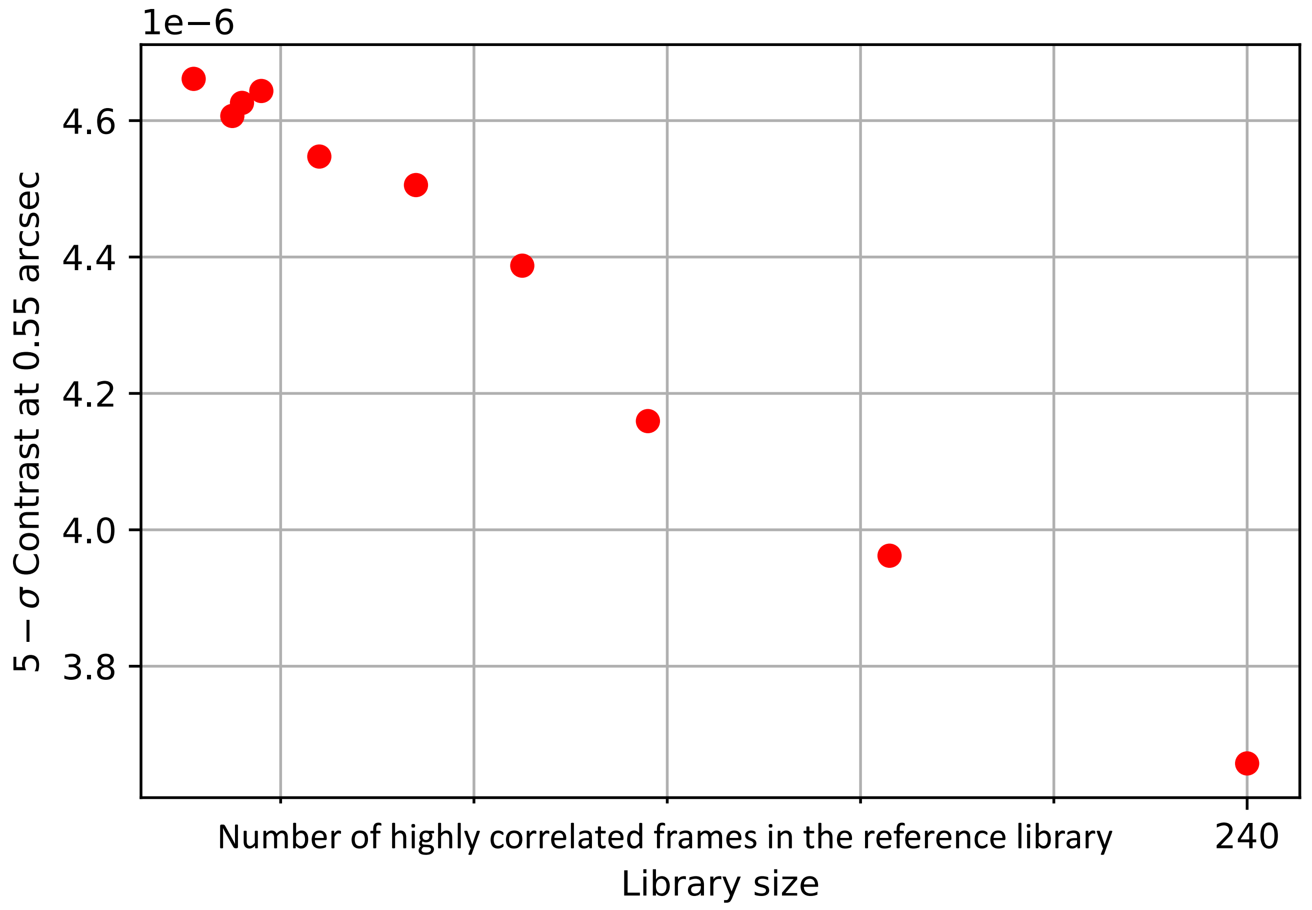


Fig. 2. *Left panel:* Contrast curves (5σ) for the selected sample with the RDI (solid lines) or ADI (dashed lines) reduction, averaged over targets in each bin of seeing conditions. *Right panel:* RDI contrast improvement over ADI. The error bar represents the scatter of the RDI improvements among targets in each bin. Four bins of seeing conditions were used.

The RDI gain is moderate: clear gain in the largest seeing bin (0.4-0.6”) below 0.4 arcsec

How to improve the contrast? Impact of the library size

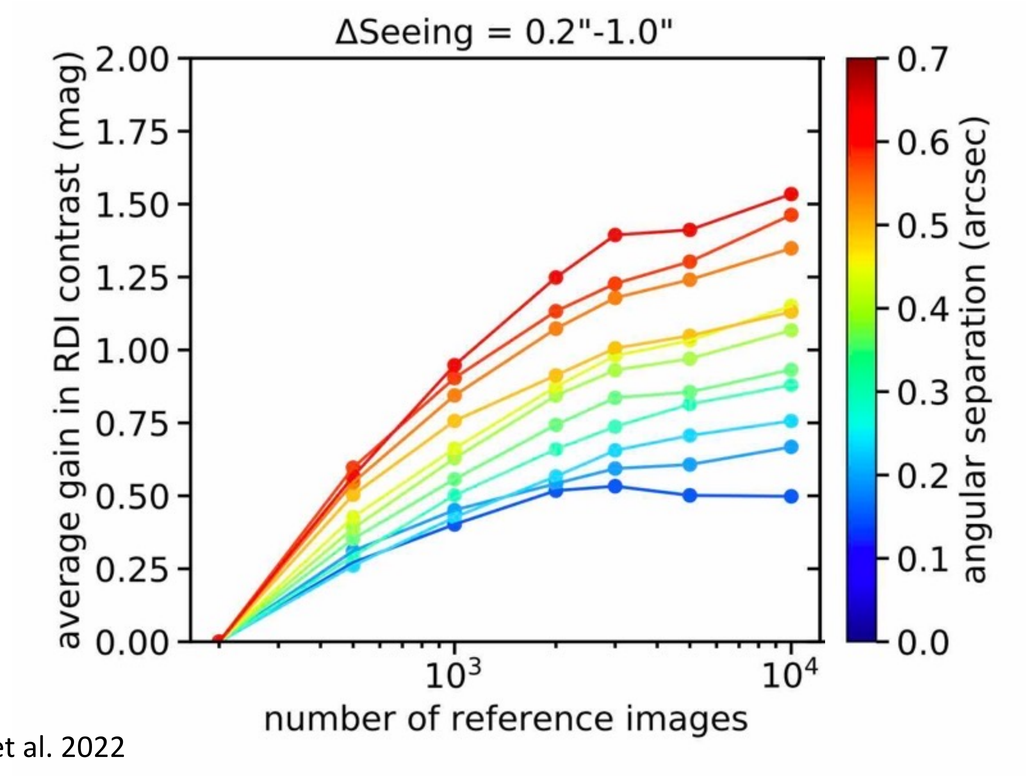


Results from the SHARDDS survey

Larger libraries of frames highly correlated with the science target improve the contrast

How to improve the contrast?
Impact of the library size

Similar conclusion found in Xie et al. 2022. Plateau a ~3000-5000 frames (4-7% of the master reference library)



Xie et al. 2022

Larger libraries of frames highly correlated with the science target improve the contrast

How to improve the contrast?
Impact of the master library size

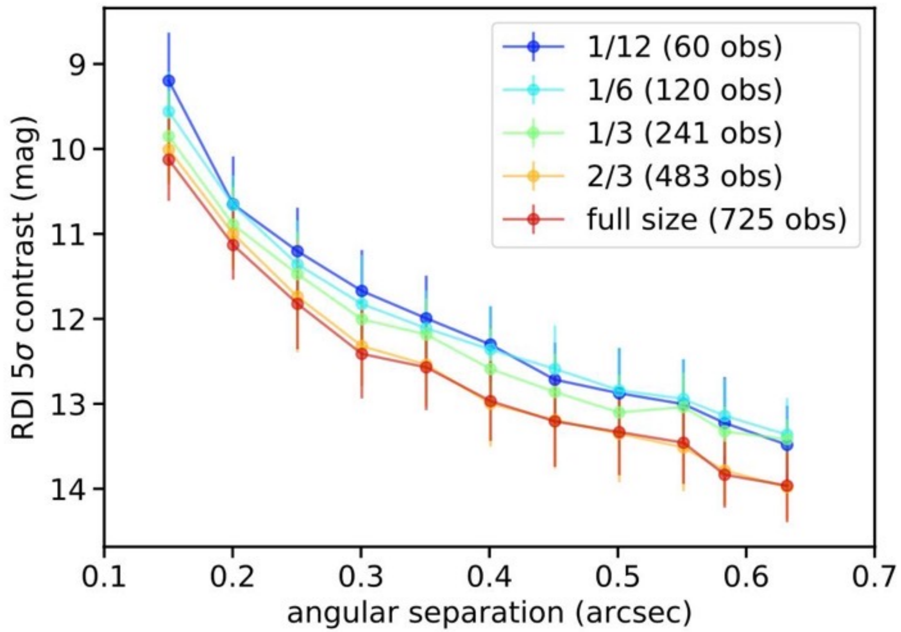


Fig. 6. Contrast curves in 5σ obtained with RDI using different sizes of master reference library. The fraction and number of observations contained in the master reference library is shown in Figure. A gain of 1 mag at 0.15″ separation can be achieved by increasing the size of master reference library from 60 observations (1/12) to 725 observations (full size).

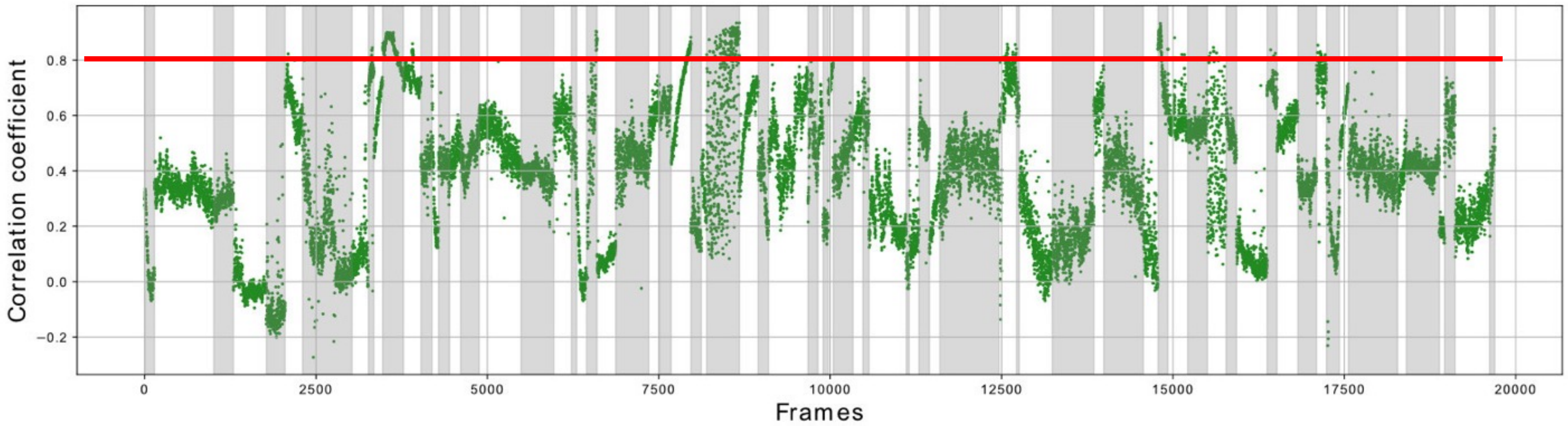
The larger the master library, the deeper the contrast

How to select highly correlated frames?

Various metrics can be used to build a similarity matrix

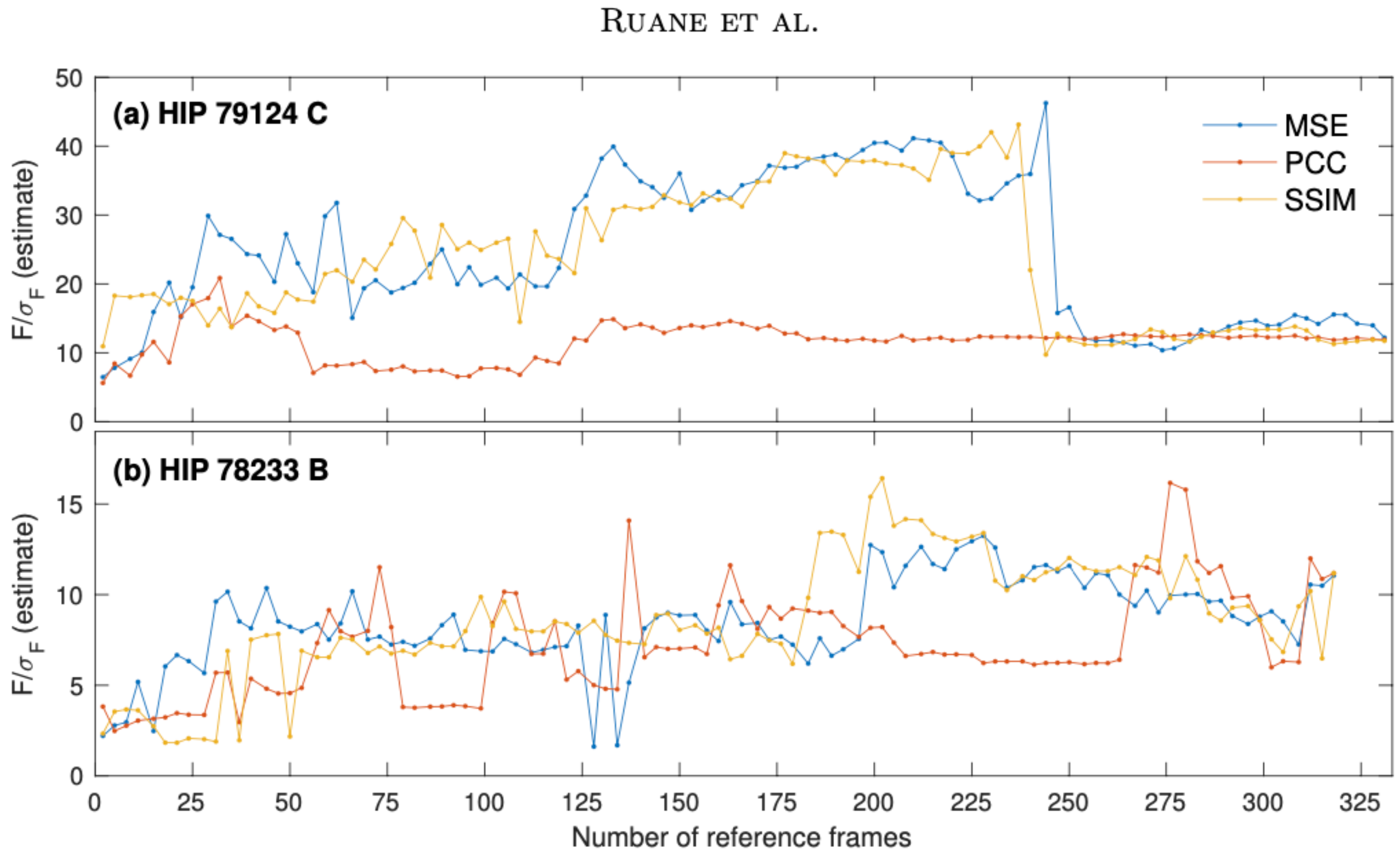
Potential reference frames

Science frames



- Pearson correlation coefficient or mean square error (MSE)
- Computer vision tools : SSIM (Structural Similarity Index Metric)
- AI tool: descriptor (next slide)
- Use of instrument, telescope, AO telemetry and environmental data (next slide)

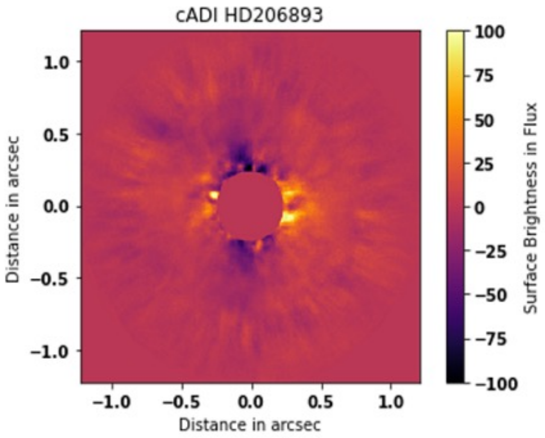
Comparison of various metrics for image similarity

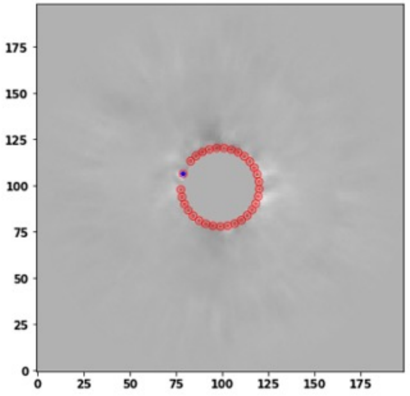


Ruane et al. 2018 found MSE and SSIM yield the best S/N on known companions

Comparison of various metrics for image similarity (HD206893)

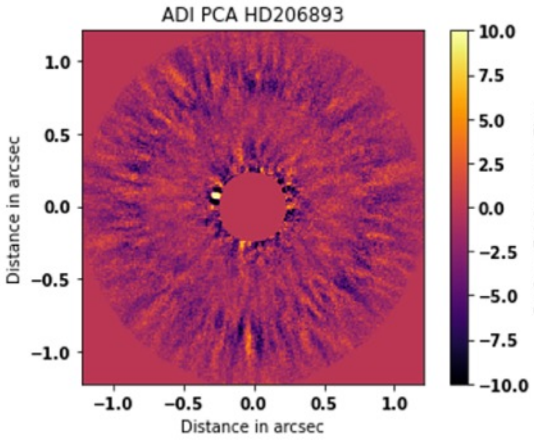
Classical ADI

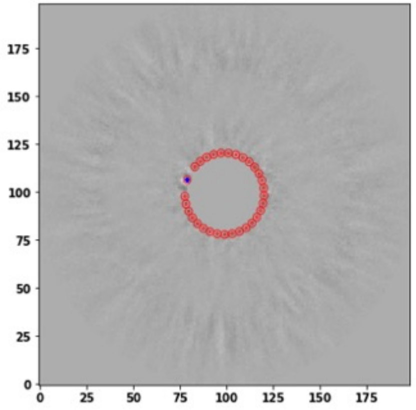




S/N 2.2

ADI PCA

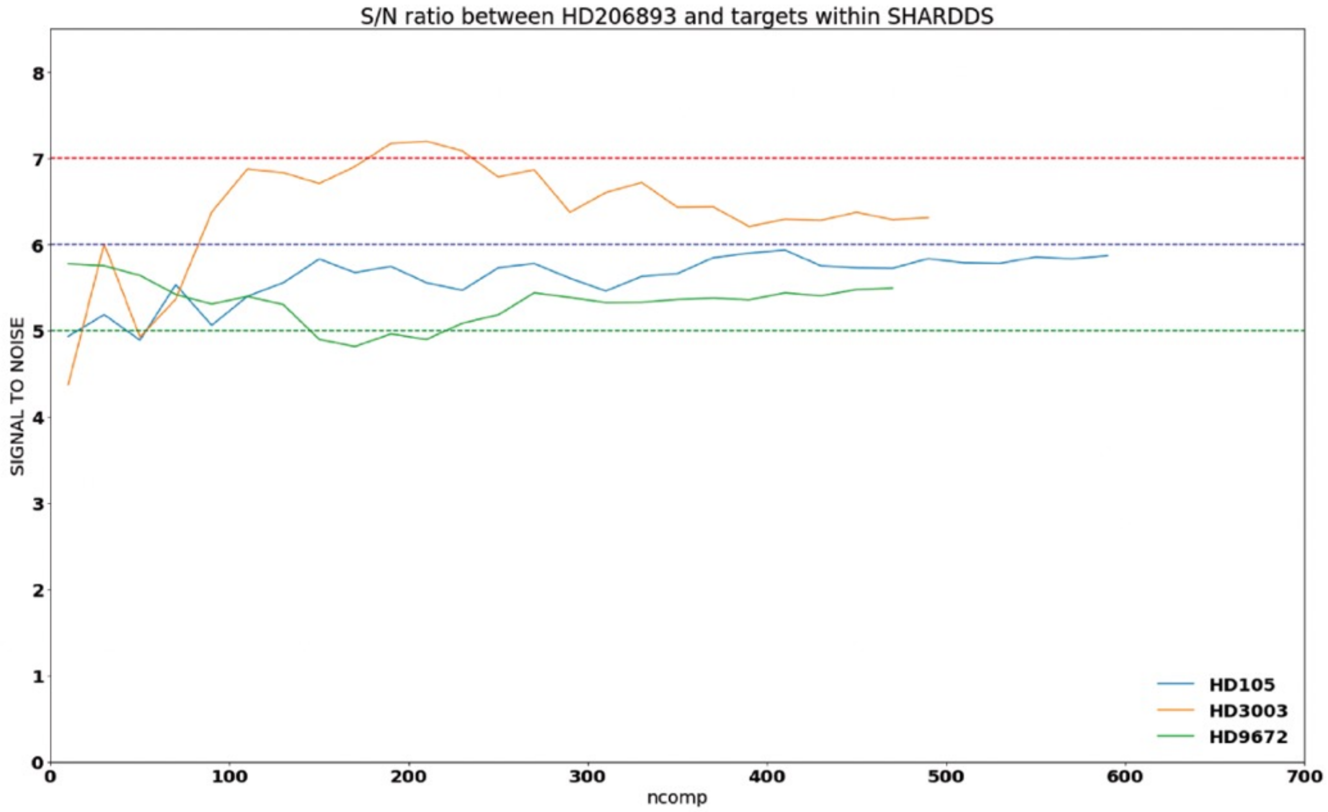




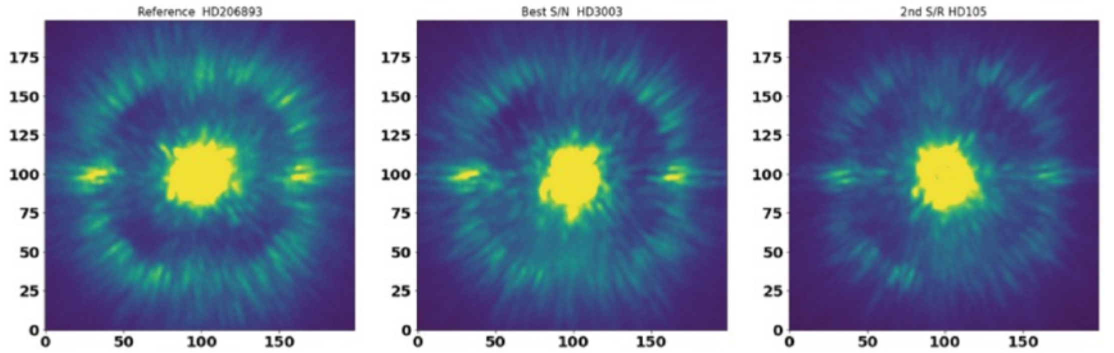
S/N 7.3 ncomp :40

S/N (VIP; Gomez Gonzalez et al. 2017)

Comparison of various metrics for image similarity (HD206893)



The graph shows the evolution of the signal to noise ratio vs. the number of components used, where the best signal is observed when RDI is performed with HD3003 as reference star.



ncomp 210
librery HD3003
S/N 7.2

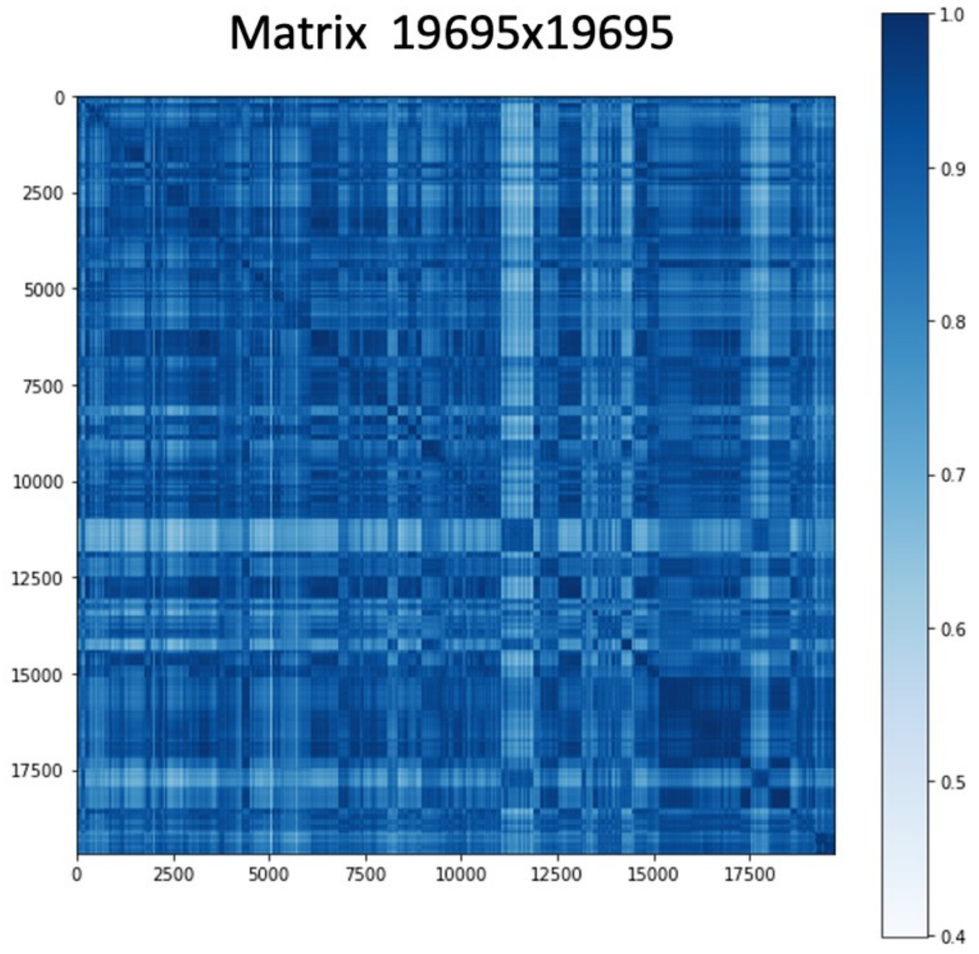
ncomp 400
llbrery HD105
S/N 5.9

After performing RDI between all SHARDDS objects and our science target we obtain that the best signal to noise ratio is 7.2 and this is after using HD 3003 as reference star.

RSDI with one target from the dataset has a poor performance compared with ADI PCA.

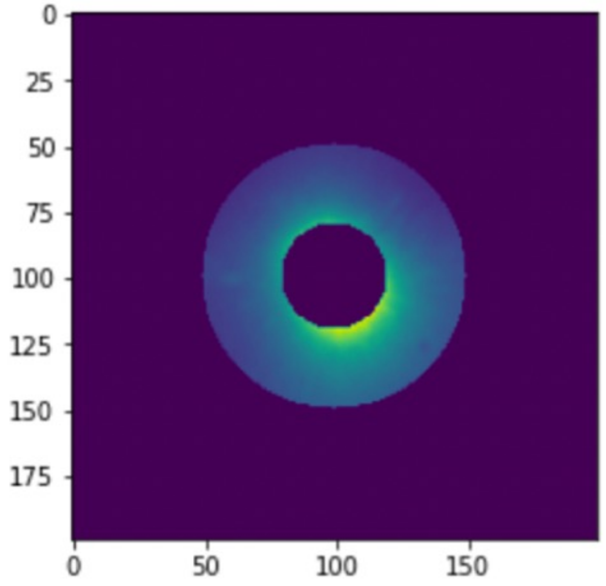
Comparison of various metrics for image similarity (HD206893)

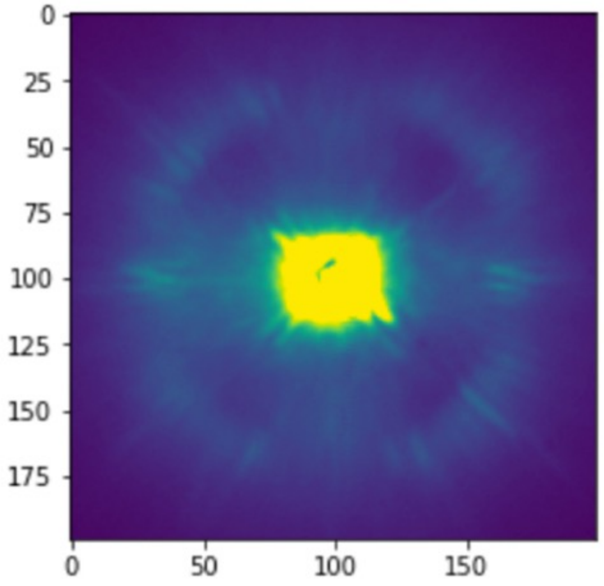
- 19695 images
- Methods :
  - Cosine Distance.
  - Similarity image index.
  - Correlation.
  - Strehl.
  - Closest per frame.
  - Random.



Comparison of various metrics for image similarity (HD206893)

Images with mask and full frame used to created the libraries.





In order to verify if there is any difference between executing the procedure with a specific zone of the images, a mask ranging from 20 to 50 pixels in radius was created.

After that, the different matrices were re-calculated and the construction criteria of the reference cube was applied again in order to create a new library with this mask already incorporated.

Comparison of various metrics for image similarity (HD206893)

- 19695 images
- Methods :

Method	Full frame S/N	Mask S/N
Cosine Distance.	8.3	6.2
Similarity image index.	7.0	8.6
Correlation.	8.2	9.1
Strehl.	7.2	7.2
Closest per frame.	7.6	7.6
Random.	8.8	8.8

Comparison of various metrics for image similarity (HD206893)

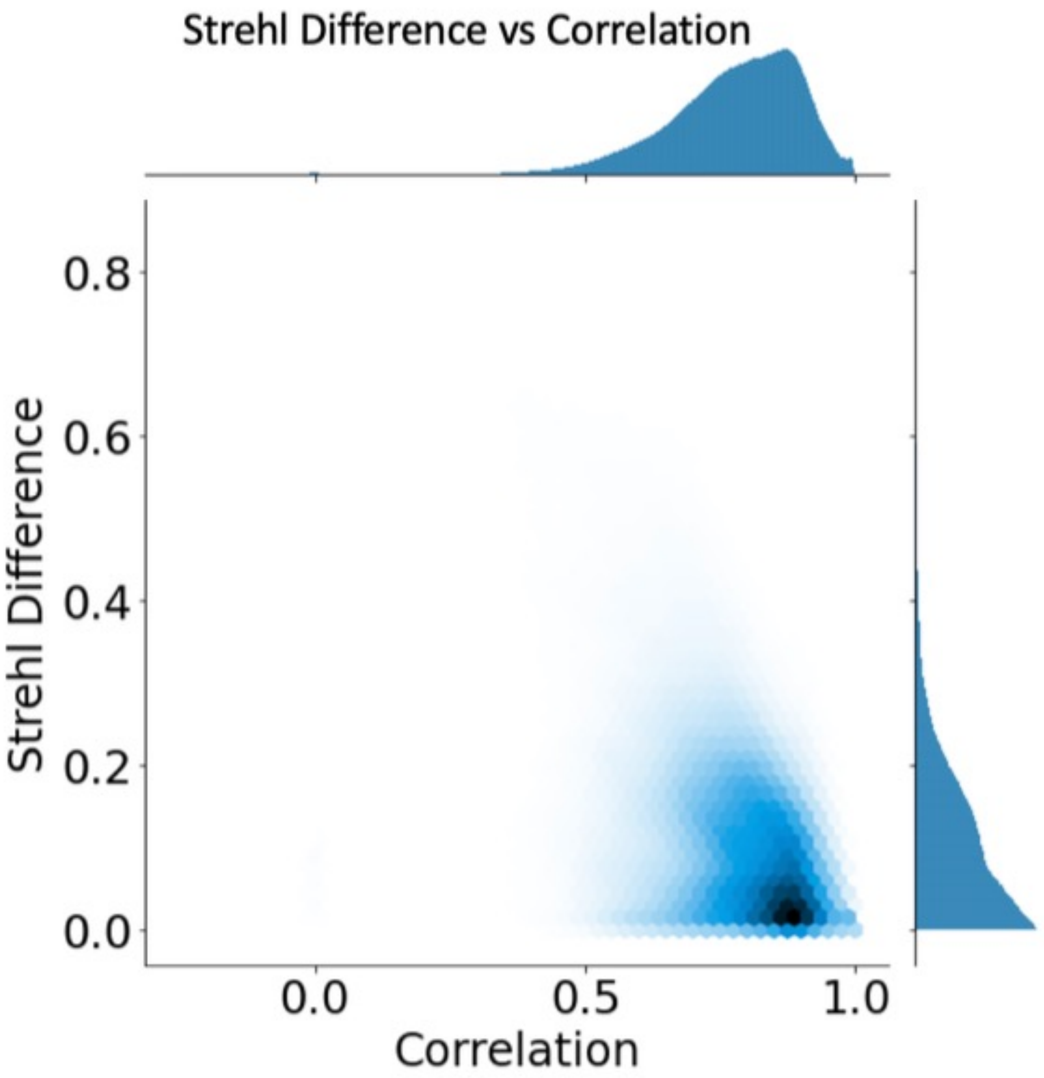
- The problem of RDI vs ADI is a multi-parameter problem where both the companion contrast, the range of parallactic angle variation and the separation of the companion come into play, without mentioning the suitability of the reference library.

- ***Reducing HD 206893 with RSDI using each target of our data set as reference, has shown the best signal-to-noise value is 7.2 with the target HD 3003. Given this result, ADI PCA delivers better results than RSDI using SHARDDS objects without library optimization.***

- According to our findings, the Pearson correlation is the image comparison method with the best overall performance. The average signal-to-noise ***value of this method within the top 10 results is 8.01***, better that all the results in RSDI with the SHARDDS data set and also better than cADI and ADI PCA. The reference library with the best (full-frame) result, a signal-to-noise ratio of 8.2 and a contrast of $2.77 \times 10{-}5$, was created using 70 principal components and the 50 closest frames at least 10 times repeated, having a total of 601 frames. When applying the mask, we obtain an average ***signal-to-noise of 8.42 and a maximum of 9.1*** with the library of the 50 closest frames at least 40 times repeated, using a total of 233 frames.

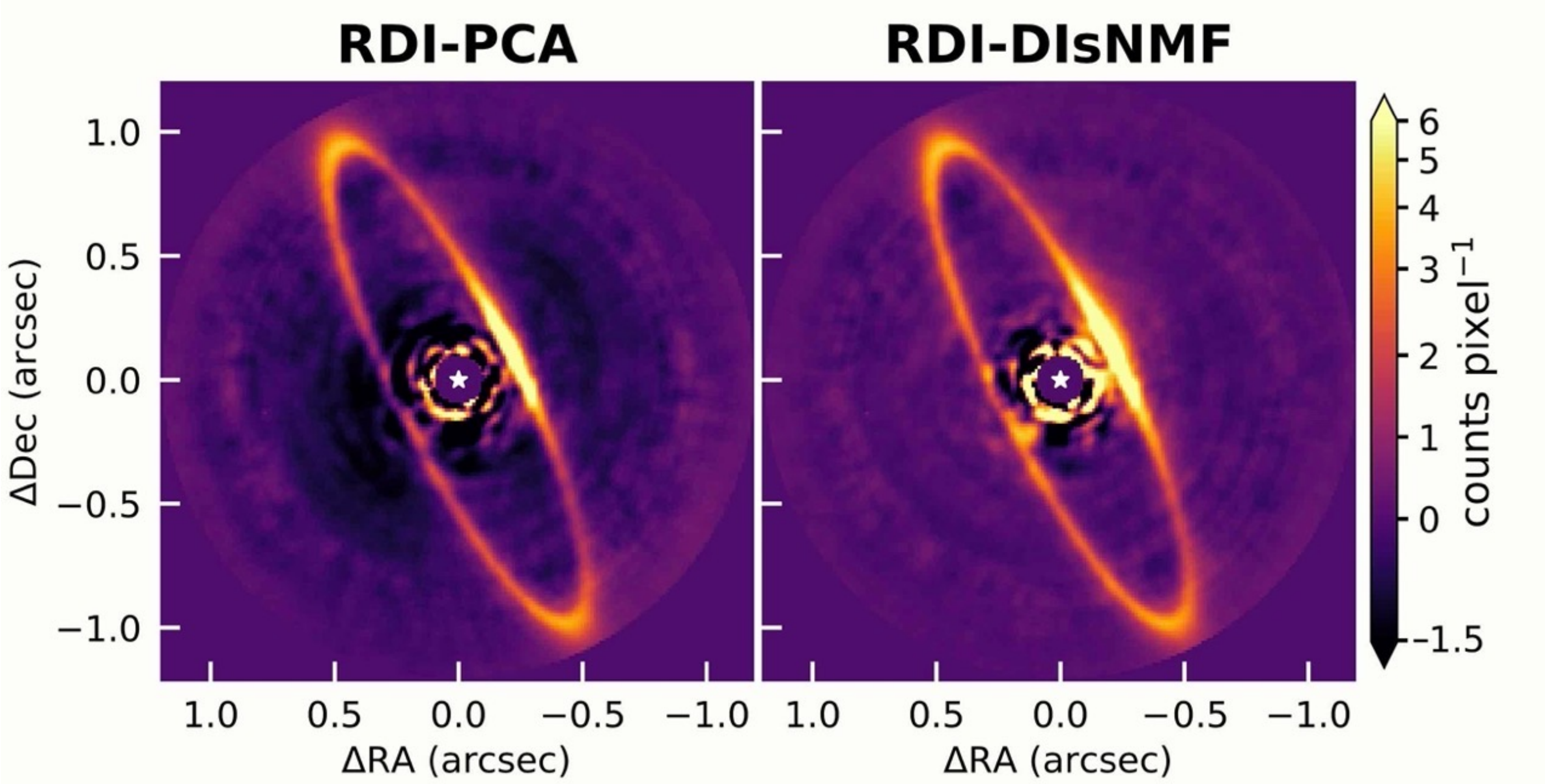
- Despite these results, the possibility of improving the performance of RSDI using artificial intelligence should not be excluded. A possible approach might be to train a CNN specifically with SPHERE data to find out if this improves the image selection.

Comparison of various metrics for image similarity



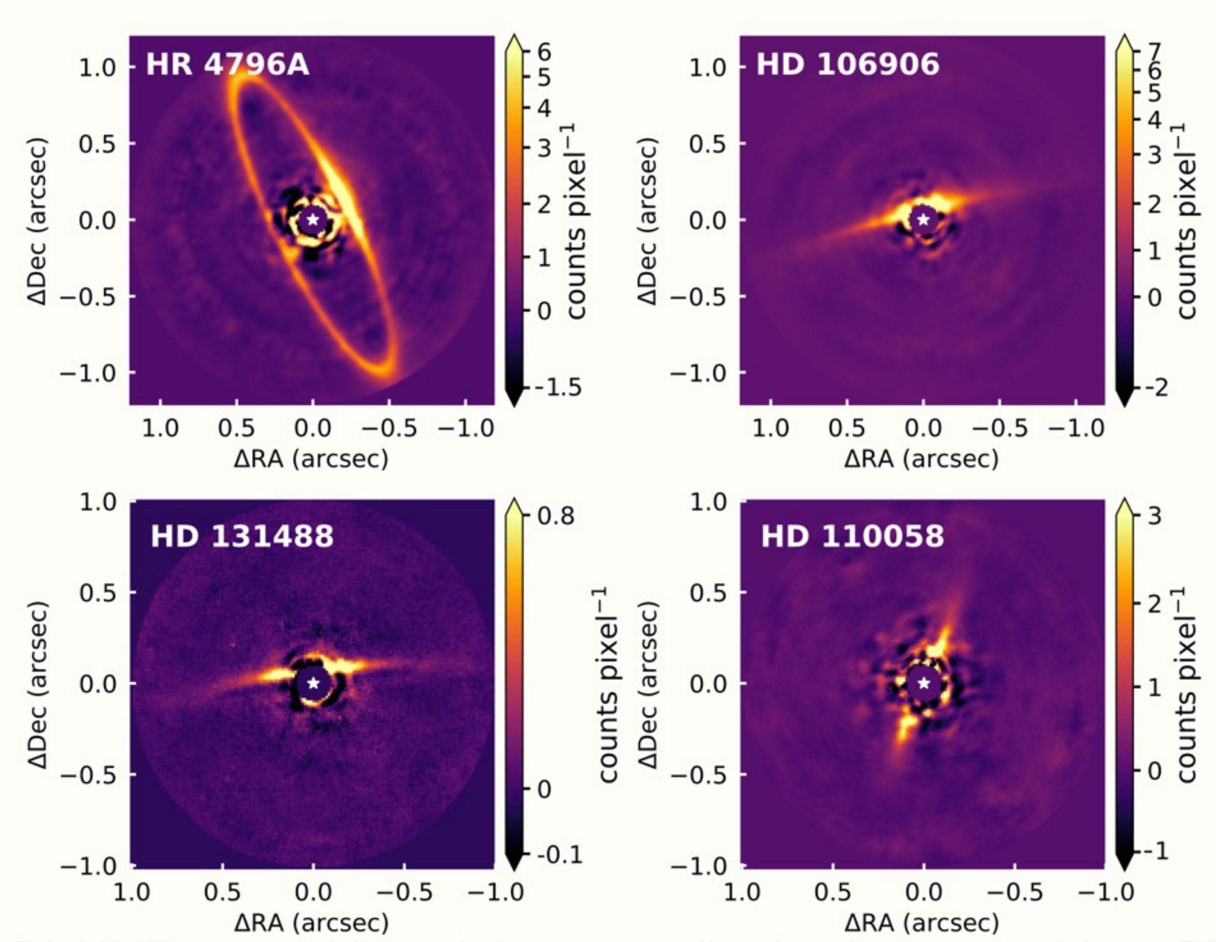
Method		contrast	S/N
cADI			2.2
ADI-PCA	Full frame	3.74×10^{-5}	7.3
	Mask	2.77×10^{-5}	8.0
RSDI-PCA	Full frame	3.93×10^{-5}	7.2
	Mask	4.22×10^{-5}	6.8
Closest per frame (R >1, N >1)			
CNN	Full frame	3.73×10^{-5}	7.1
	Mask	3.96×10^{-5}	6.4
Correlation	Full frame	2.61×10^{-5}	8.0
	Mask	2.58×10^{-5}	8.4
SIM	Full frame	3.21×10^{-5}	6.5
	Mask	2.62×10^{-5}	8.0
Strehl	Full frame	2.90×10^{-5}	6.5
	Mask	3.02×10^{-5}	6.1
Closest per frame (R = 1, N = 1)			
	Full frame	3.06×10^{-5}	6.9
	Mask	2.57×10^{-5}	7.0
Random			
	Full frame	2.67×10^{-5}	8.1
	Mask	2.80×10^{-5}	8.1

How to best reconstruct the PSF from a library



PCA was the standard so far, as a simple linear technique to optimally combine the references.
DIsNMF: data imputation using sequential non-negative matrix factorization (Ren et al. 2020) is promising too (Xie et al. in prep)

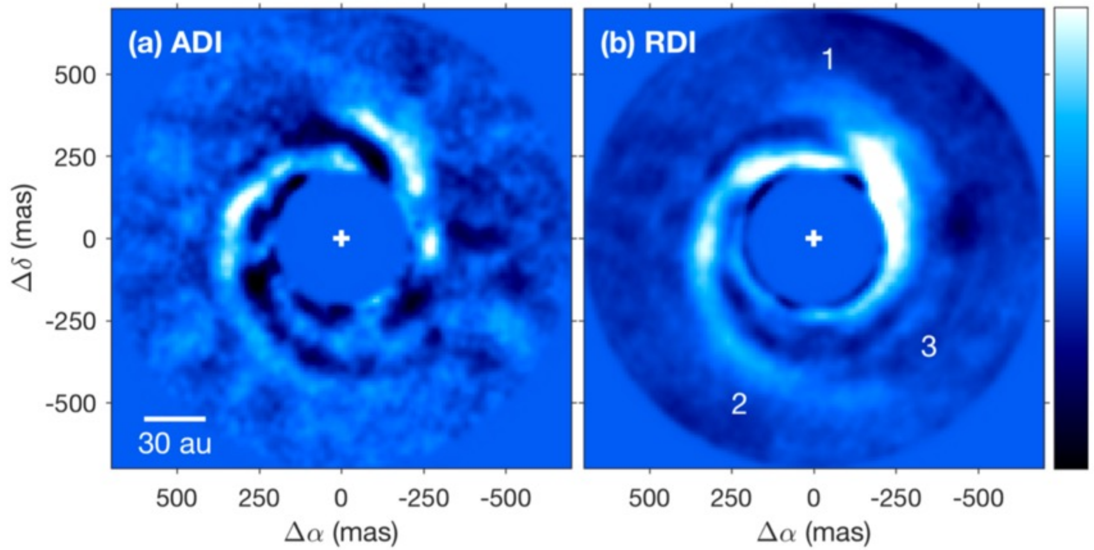
How to best reconstruct the PSF from a library



DIsNMF images of known debris disks (Xie et al. in prep).

Conclusions

- RDI applicability on VLT/SPHERE:
  - Snapshot imaging
  - Below 200 – 400 mas for planets
  - Adapted for extended objects



Ruane et al. 2019: RDI with NIRC2 and Vortex coronagraph

- Still a lot of work to fully unleash the potential of large libraries with ground-based high-contrast imagers:
  - How to best select the frames to be used as references ?
  - How to optimally combine them to subtract the halo ?

- How to provide such libraries to astronomers ?

  Clever archive system allowing customized queries → Sophia's PhD project